

Basado en las investigaciones teóricas sobre el estado de la técnica, se establecieron etapas funcionales para la implementación de un sistema de IoT.

Algunos de los resultados de esta investigación son los siguientes:

- El diseño de un circuito electrónico, junto con su *firmware* específico, capaz de realizar la lectura del carné RFID de cada estudiante que ingresa al aula o recinto; y de enviar estos datos a internet vía conexión Wi-Fi.
- La creación de un código *script* dentro de los servicios del Google App Script, programa que corre en internet y que está pendiente de recibir la información generada por el circuito electrónico cada vez que un estudiante escanea su carné al entrar.
- Aplicación el sitio de Google, junto con la configuración de hojas de cálculo de Google Drive, para la visualización en tiempo real de la información del carné RFID escaneado.

BREVE HOJA DE VIDA DE LOS INVESTIGADORES

Ronny Adalberto Cortez Reyes. Actualmente es investigador a tiempo completo en el área de Tecnología y asistente de docente en el área de Redes en la Utec.

Omar Otoniel Flores Cortez. Investigador y docente titular del departamento de Electrónica de la Utec, en las áreas de Programación de plataformas para sistemas embebidos y de Robótica y domótica educativa.

Verónica Idalia Rosa de Rivera. Docente investigadora de la Utec. Actualmente, candidata a doctora en Informática de la Universidad de Alicante, España. Máster en *Visual Analytics* y *Big Data* (2014) de la Universidad de La Rioja, España.

Autoridades Utec

Dr. José Mauricio Loucel
Presidente

Lic. Carlos Reynaldo López Nuila
Vicepresidente

Ing. Nelson Zárate
Rector Utec

Licda. Noris Isabel López Guevara
Vicerrectora de Investigación y Proyección Social

Dra. Camila Calles Minero
Directora de Investigaciones

OTRAS LÍNEAS DE INVESTIGACIÓN UTEC

- Turismo
- Democracia y gobernabilidad
- Vivienda y desarrollo urbano
- Desarrollo e innovación tecnológica

INVESTIGACIÓN EN BREVE

Es una colección de fascículos que resumen los resultados de las investigaciones realizadas por la Vicerrectoría de Investigación y Proyección Social.

No hay enseñanza sin investigación ni investigación sin enseñanza
Pablo Freire

VICERRECTORÍA DE INVESTIGACIÓN Y PROYECCIÓN SOCIAL

Calle Arce y 19ª avenida Sur n.º 1045, edificio *Dr. José Adolfo Araujo Romagoza*.
San Salvador, El Salvador, (503) 2275-1013 / 2275-1011



www.utec.edu.sv

Centro de llamadas: 2275-8888
Maestrías: 2275 2700



¡¡AGAMOS LA DIFERENCIA

**Universidad Tecnológica
de El Salvador**



n.º 17

AGOSTO 2018

INVESTIGACIÓN EN BREVE

**Universidad Tecnológica
de El Salvador**



Vicerrectoría de Investigación y Proyección Social

Compilación de investigaciones de tecnología 2017

Extracción de conocimiento a partir de texto

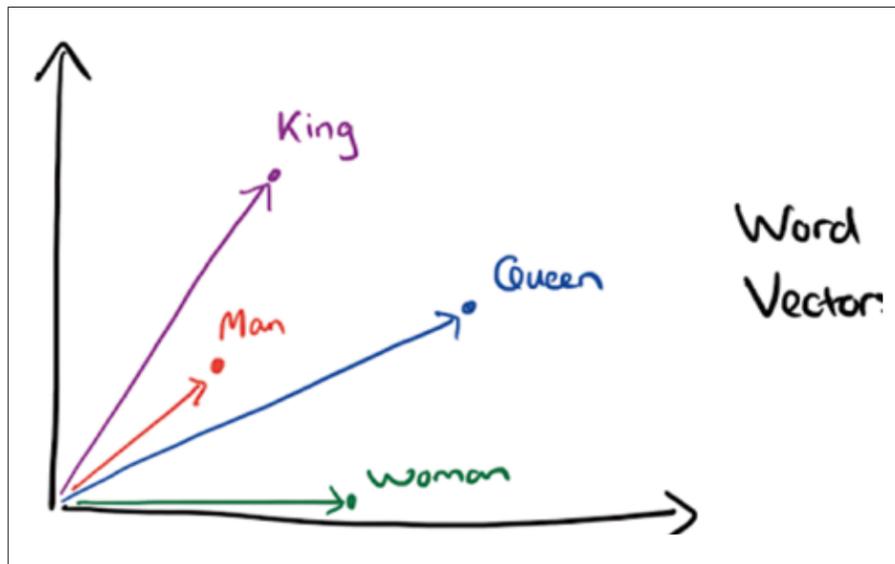
Investigador:
Ronny Adalberto Cortez Reyes

Aulas conectadas: sistema IoT para el registro de asistentes

Investigadores:
Omar Otoniel Flores Cortez
Verónica Idalia Rosa Urrutia

En esta ocasión, la Universidad Tecnológica de El Salvador presenta dos investigaciones en el área de tecnología: "Extracción de conocimiento a partir de texto" y "Aulas conectadas: sistema IoT para el registro de asistentes". Ambas muestran la utilización de herramientas tecnológicas para la obtención de resultados de investigación, además analizan la aplicación de esos resultados en dos ámbitos: las aulas y el análisis de texto.

Compilación de investigaciones de tecnología 2017 “Extracción de conocimiento a partir de textos” y “Aulas conectadas: sistema IoT para el registro de asistentes”



www.depends-on-the-definition.com
<https://www.depends-on-the-definition.com/guide-to-word-vectors-with-gensim-and-keras/>

En esta ocasión, la Universidad Tecnológica de El Salvador (Utec) presenta dos investigaciones en el área de tecnología: “Extracción de conocimiento a partir de textos” y “Aulas conectadas: sistema IoT para el registro de asistentes”.

La Utec, en su constante apuesta por la tecnología, desarrolla investigaciones en dos grandes áreas que son de relevancia en la actualidad: internet de las cosas (IoT, acrónimo del inglés) y *big data*. Como resultado de este interés, se presentan dos investigaciones: “Extracción de conocimiento a partir de textos” y “Aulas conectadas: sistema IoT para el registro de asistentes”, ambas enfocadas en la implementación de nuevas tecnologías para la solución de problemas de forma eficiente.

El trabajo “Extracción de conocimiento a partir de textos” tiene como objetivo aplicar técnicas de *data mining* para, a partir de un conjunto grande de textos, obtener información que permita extraer datos útiles para generar conceptos, ontologías y definiciones.

En los últimos años, se ha generado una gran cantidad de textos en formato digital en diferentes plataformas, como, por ejemplo,

redes sociales, correos, publicaciones científicas, foros, comentarios, periódicos, esto debido a que el número de usuarios con acceso a internet es cada vez mayor y con diferentes tipos de dispositivos; y esta tendencia puede mantenerse debido a las nuevas tecnologías.

Ante el incremento exponencial de datos en forma de texto y a la dificultad de su análisis mediante métodos tradicionales supervisados por humanos y que requieren una gran cantidad de tiempo y expertos que muchas veces son escasos, se hace necesario el conocimiento e implementación de técnicas de minería de datos que permitan la obtención de ontologías para describir los conceptos el menor tiempo posible, en forma eficiente y con bajos costos.

Por medio de técnicas de *data mining*, se puede procesar grandes cantidades de texto sin la necesidad de que un experto dedique todo su tiempo a dicho proceso, obteniendo resultados en mucho menor tiempo. Para el proyecto, se decidió utilizar Word2Vec debido a sus ventajas, ya que es eficiente de entrenar, fácilmente disponible en línea con código y modelos preentrenados.

Se han utilizado como fuente de información artículos descargados de Scopus que contienen *title*, *abstract*, *keywords* y *author keywords*, y un conjunto de noticias de noviembre 2017 formadas por un titular, resumen y cuerpo proporcionado por la Unidad Focus Data de *El Diario de Hoy* para la etapa de pruebas y aprendizaje en el uso de Word2Vec, y un conjunto de tuits de cuentas pertenecientes a algunos candidatos seleccionados basado en su trascendencia mediática, manteniendo la representatividad tanto de candidatos a diputados de la Asamblea Legislativa como a ser miembros de concejos municipales; y con un balance entre los contendientes de diferentes partidos políticos para hacer un análisis que incluya relación entre las palabras, frecuencia de aparición en los textos; representación visual, tales como *heatmap*, dendrogramas, clústeres y nubes de palabras para mostrar los resultados obtenidos.

Las pruebas realizadas fueron preprocesamiento de texto, aplicación de los modelos sobre cada una de las partes de los artículos, uso de las nubes de palabras y otros métodos gráficos para determinar si el modelo es capaz de extraer información de utilidad y su calidad. Para los textos extraídos de los tuits, se hizo una limpieza propia del contenido.

Teniendo en cuenta nuestras necesidades y el contexto de las pruebas, los mejores resultados con Word2Vec se han obtenido utilizando el modelo que se ha generado sin utilizar *stemming* sobre los encabezados de las noticias.

Con los resultados obtenidos, se ha identificado que Word2Vec es una poderosa herramienta, con la cual se puede extraer mucha información a partir de textos en los que se tiene en cuenta el contexto de cada una de las palabras, aun utilizando un modelo genérico a partir de una inmensa cantidad de textos que puede abarcar una variedad de términos, palabras y situaciones.

El análisis de textos provenientes de redes sociales, como es el caso de los tuits, se vuelve

aún más complicado que el de los textos originados por otras fuentes con una estructura más rígida o que pasan por un filtro antes de ser publicados, debido a la libertad con que las personas los escriben, como, por ejemplo, usando abreviaturas, incorrecciones, términos propios de cada país e incluyendo caracteres especiales, símbolos, URL y entidades de *retweet*, entre otros.

Mediante los tuits, se pudo extraer información de los candidatos analizados, de sus propuestas e ideas; algunos de ellos con mayor movimiento en las redes sociales, tanto en cantidad como en el número de palabras utilizadas. Entre los usos de las cuentas, estas fueron informativas, de propuestas y de noticias.

AULAS CONECTADAS: SISTEMA IOT PARA EL REGISTRO DE ASISTENTES

En el caso de la investigación aplicada “Aulas conectadas: aplicación del IoT para el registro de asistentes”, desarrollada durante 2017, se propone nuevo conocimiento científico dentro de la aplicación de tecnologías y técnicas del IoT en la solución de una situación problemática específica: el registro automatizado de los asistentes a un evento, recinto, aula, conferencia, etc.; y cuyo objetivo principal fue el diseño, desarrollo y validación de una plataforma o sistema IoT de bajo costo para el registro automatizado, vía internet, de datos de asistentes a un recinto basados en la información recogida por sensores electrónicos inalámbricos dispuestos en la entrada del lugar; y que sea una herramienta de apoyo a las labores de registro de datos de usuarios y para el fortalecimiento de la seguridad de las personas que ingresan a la institución.

Actualmente, el uso de sistemas electrónicos automatizados aplicados en la solución de tareas cotidianas o repetitivas es una necesidad; optimizar procesos y recursos económicos y humanos es prioritario para toda institución moderna. En el caso de estudio específico, de la Universidad Tecnológica de El Salvador, uno de los procesos de índole administrativo que más demanda tiempo y esfuerzo, por parte del cuerpo académico de cada escuela y decanato, es el conteo de los estudiantes asistentes a las sesiones de clases en los diferentes horarios. Este proceso de recolección de datos se está desarrollando de

forma manual; y en ocasiones hasta su procesamiento y notificación.

El registro de asistencia en un centro educativo permite llevar un mayor control de la población estudiantil, con lo que se puede sacar ciertos estadísticos para la toma de decisiones a corto y largo plazo. La realidad salvadoreña, cada día, demanda una mayor atención en el área de la seguridad, por ejemplo, una institución como la Utec, que es la universidad privada de El Salvador con la mayor población estudiantil, se ve en la necesidad de depender de un mecanismo que proporcione control de quienes ingresan a la institución, y además, que sea un claro mensaje para las personas mal intencionadas ajenas a la institución que el alto nivel de monitoreo es permanente, por lo que se propone un sistema de bajo costo que permita tener en cada aula un dispositivo electrónico que registre a cada estudiante que ingresa a un aula mediante el uso de un carné RFID, similar a los utilizados por el cuerpo docente hora-clase; y que además notifique en tiempo real, vía acceso a un sitio web, de la cantidad de asistentes a clases a los administradores, directores y decanos responsables, así como el listado de la identificación de quienes han asistido.

El objetivo general de la investigación consiste en implementar un sistema IoT de bajo costo para el registro automatizado, vía internet, de datos de asistentes a un lugar, como apoyo a las labores de registro de datos y fortalecimiento de la seguridad de las personas que ingresan a la institución. Los objetivos específicos son los siguientes:

- Aportar y divulgar nuevo conocimiento científico, teórico y práctico, sobre el diseño e implementación de sistemas de IoT eficientes y de bajo costo en la solución de problemas de automatización.
- Diseñar y construir un circuito electrónico basado en un sistema microcontrolador que sea capaz de registrar datos de un usuario que utilice tarjetas RFID para su ingreso a un evento/clase o lugar específico, además de permitir enviar estos datos, cantidad y nombres, por medio de una red local inalámbrica.
- Diseñar un *firmware* exclusivo para el funcionamiento del prototipo electrónico, escrito en lenguaje C++, que sea capaz de controlar el prototipo propuesto.
- Construir y acoplar el prototipo diseñado en un aula para evaluar resultados de producción.

La investigación fue desarrollada usando el método multimodal-experimentación sin hipótesis, el cual es, según Roberto Hernández Sampieri (1991), “un método empleado en casos donde la investigación tiene por objeto el provocar determinados fenómenos que no se presentan usualmente en la naturaleza y cuyo conocimiento puede ser interesante o importante en el avance de la ciencia y la tecnología”. Para la verificación del diseño propuesto, se realizó la implementación del diseño fruto de esta investigación aplicada dentro de un ambiente académico real: el campus central de la Utec.



squarespace.com